An Examination of How Teachers' Beliefs about Scientific Argumentation are Impacted by Multimedia Educative Curriculum Materials (MECMs)

Katherine L. McNeill, María González-Howard, Lisa M. Marco-Bujosa, Suzanna Loper & Laura O'Dwyer

Boston College

contact info:
Katherine L. McNeill
Lynch School of Education, Boston College
140 Commonwealth Avenue, Chestnut Hill, MA 02467
Phone: 617-552-4229
Fax: 617-552-1840
kmcneill@bc.edu

Reference as:

McNeill, K. L., González-Howard, M., Marco-Bujosa, L., Loper, S., & O'Dwyer, L. (2017, April). *An examination of how teachers' beliefs about scientific argumentation are impacted by multimedia educative curriculum materials (MECMs).* Paper to be presented at the annual meeting of the National Association for Research in Science Teaching (NARST), San Antonio, TX.

An Examination of How Teachers' Beliefs about Scientific Argumentation are Impacted by Multimedia Educative Curriculum Materials (MECMs)

Abstract

Recent reform efforts include a shift to focusing on science practices. Teachers require support in integrating these science practices into their classroom instruction. Multimedia educative curriculum materials (MECMs), which are digital materials explicitly designed to support teachers, offer one potential resource for this critical need. Consequently, we investigated how teachers used MECMs and whether that use impacted teachers' beliefs about scientific argumentation. We conducted a randomized experimental study with 90 middle school science teachers. Both control and experimental groups taught the same curriculum, using a web-based teacher's guide. Additionally, the experimental teachers received the MECMs including 24 videos and 17 interactive reflective prompts. We collected multiple data sources: pre surveys, backend curriculum use, self-report curriculum use, and post surveys. The results suggest that enacting a curriculum with a focus on argumentation can support positive changes in teachers' beliefs about argumentation. Furthermore, we observed a wide range in how teachers used the curriculum. In terms of self-efficacy, this differential use impacted teacher's beliefs. Teachers became more confident in their ability to teach argumentation as they enacted more lessons. Additionally, the MECM teachers had lower changes in self-efficacy, perhaps because the video may have problematized what teachers thought counted as argumentation.

An Examination of How Teachers' Beliefs about Scientific Argumentation are Impacted by Multimedia Educative Curriculum Materials (MECMs)

Realizing the vision set forth by recent standards (NGSS Lead States, 2013) requires a shift in how science instruction has been typically carried out. For teachers, a critical aspect of this change will be supporting their students in science practices (Pruitt, 2014). Educative curriculum materials (Davis & Krajcik, 2005) hold promise for helping teachers around the science practices. Much of the work around educative curriculum has focused on print-based materials, which are unable to capture the complex discursive interactions inherent to many science practices (Loper, McNeill, & González-Howard, 2017). Technology, however, affords the creation of multimedia educative curriculum materials (MECMs), such as classroom videos, which may better support teachers around science practices. Furthermore, the beliefs teachers hold influence the decisions they make about using curriculum (Bryan, 2012), and so it is important to consider the ways educative materials might address teacher beliefs. Thus, in this study we examine the impact of MECMs on teacher beliefs, specifically for the practice of scientific argumentation. Specifically, we investigated the following research questions:

1. Did teacher's beliefs for scientific argumentation change from pre to post?
2. What variation existed in teachers' use of the digital curriculum and MECM supports?
3. What impact did the MECMs and teachers' use of the curriculum have on teachers' beliefs?

Conceptual Framework

Scientific Argumentation

Consistent with other researchers, we focus on two interrelated aspects of scientific argumentation: structure and dialogic interactions (Jiménez-Aleixandre & Erduran, 2008; McNeill, Gonzalez-Howard, Katsh-Singer, & Loper, 2016). Argument structure consists of a claim justified by evidence and reasoning (McNeill et al., 2006). The dialogic components of argumentation focus on how students engage with their peers to construct, critique, and revise arguments (Ford, 2012). Dialogic interactions differ greatly from the typical classroom discourse, where students generally interact with the teacher rather than other students (Berland & Reiser, 2011). Despite the recent push for argumentation, students rarely have opportunities to do this in science classrooms (Osborne, 2010).

The role of the teacher is essential for supporting students in argumentation (Evagorou & Dillon, 2011). For example, in a study of curriculum supports for leading discussions in high school science, Alozie, Moje, and Krajcik (2010) found that teachers tended to rely upon traditional recitation formats for classroom discussions. They concluded that curricular supports were necessary to help teachers promote dialogic interactions in their classroom. Additionally, McNeill and colleagues (2016) found that teachers tend to focus on the superficial aspects of argumentation, such as the language of an argument (e.g., claim, evidence, reasoning) without deeper understandings of the practice of argumentation. They referred to this practice as "pseudoargumentation" because teachers do not engage students in the practice of argumentation. The authors suggest helping teachers focus more on the quality of dialogic interactions between students with respect to how students build off of and critique each other's claims could be productive in supporting teacher PCK of argumentation. These studies indicate teachers may struggle with their own understanding and instructional practices supporting student learning of argumentation.

Teachers Beliefs and Scientific Argumentation

The struggles teachers face in their knowledge and instruction of argumentation may be due to teachers holding differing beliefs about argumentation. Research indicates teacher beliefs significantly impact the instructional decisions they make. This includes teacher beliefs about science and beliefs about argumentation's role in science education, beliefs about science teaching, and varied perceptions of which students can engage in argument (Sampson & Blanchard, 2012).

Science teachers hold personal epistemologies about science, or beliefs about knowledge and the nature of knowledge. Research indicates that how teachers view and understand science as a way of knowing may influence their students' beliefs. Argumentation is fundamentally epistemological as its focus is on generating knowledge by constructing an argument through the use of evidence and reasoning. Teachers' justifications of their own knowledge of science (handed down by authority vs. inductively derived through reasoning) is essential in their curricular decisions and influencing students' beliefs (Jones & Leagon, 2014). Similarly, through surveys and interviews with middle school science teachers, McNeill, Katsh-Singer, Gonzalez-Howard, and Loper (2016) found teachers' own science learning goals impacted their argumentation instruction. With respect to beliefs about teaching, Zohar (2008) found that science teachers who held teaching beliefs focused on the transmission of knowledge prioritized providing factual information and correct answers to students in their science instruction. In contrast, teachers whose beliefs more closely aligned with the idea of classroom 'knowledge construction' engaged students in more problem-solving and critical thinking opportunities.

Similar results have been found in the little research that exists around teacher beliefs and scientific argumentation; teacher beliefs do impact argumentation instruction (Sampson & Blanchard, 2012). Teachers' beliefs about their students' abilities to engage in argumentation may also impact their instructional decisions (Katsh-Singer, McNeill, & Loper, 2016; Sampson & Blanchard, 2012). For example, Katsh-Singer and colleagues (2016) found that teacher beliefs about student ability to engage in argumentation vary based upon factors such as the socioeconomic status of their students (Katsh-Singer, McNeill & Loper, 2016). These kinds of beliefs may cause teachers to undermine the goals of argumentation by placing an instructional priority on transmitting knowledge. Educative curriculum materials have potential for attending to teachers' beliefs around argumentation.

Educative Curriculum Materials

When intentionally created to align with reform efforts, curriculum materials can be a concrete means by which to support teachers in making desired instructional changes (Remillard, 2005). Not only can curriculum materials help increase teachers' understanding of a particular area of instruction (e.g., content matter), but ideally they can also enable teachers to develop knowledge that can be applied to new situations (Davis & Krajcik, 2005). In terms of the complex, three-dimensional learning described by the NGSS, curriculum materials are promising for helping teachers learn how to integrate science practices into their classroom instruction (Krajcik & Delen, 2017). In particular, educative curriculum materials include features intended to support teachers, and not just students (Ball & Cohen, 1996; Davis & Krajcik, 2005). These educative features can take on many different forms, including notes in the margins of the teacher's guide (Davis, Janssen & Van Driel, 2016). Furthermore, because educative features are embedded within lessons, they provide teachers with scaffolded, authentic assistance for learning new skills and practices.

Recent work has explored how teachers use educative curriculum. For instance, Arias and colleagues (2016) examined how elementary teachers used educative features focused on the science practices. Similar to others (e.g., McNeill, 2009; Schneider, Krajcik & Blumenfeld, 2005), they saw variation in how teachers used particular curricular features. Their findings also indicated that teachers drew more from the educative features that contained classroom representations, such as narratives describing instructional challenges and rationales behind subsequent teaching moves. In another study, researchers analyzed how variation in curriculum use impacted teacher learning and implementation of the science practice of argumentation (Marco-Bujosa, McNeill, González-Howard & Loper, 2017). Marco-Bujosa and colleagues (2017) found that how teachers perceived, and consequently used, curriculum – as either a resource for supporting only student learning, or as a resource that also encompasses learning benefits for educators – influenced their learning gains around argumentation. Taken together, this work points to the different ways that teachers use, and learn from, curriculum. Yet, most of this prior research has examined text-based curriculum, and thus, research needs to begin examining how teachers utilize and learn from digital educative curriculum.

Using Technology to Support Teacher Learning

Given technological advances, curriculum materials can now be digital and accessed through web-based platforms. However, little research has explored teachers' use of digital curriculum. For instance, a study focused on a web-based curriculum that supported project-based learning, examined the ways that teachers utilized educative features of the curriculum (Duncan, El-Moslimany, McDonnell & Lichtenwalner, 2011). These researchers found that although teachers reported finding the educative features useful, they did not use them in "just in time" fashion as the developers envisioned, but instead while preparing for lessons. Moreover, it is worth noting that although the curriculum was web-based the educative features were textual, comprising of notes highlighting certain areas like student difficulties and related activities.

In terms of educative features within web-based platforms, technological affordances allow for classroom representations to be enhanced to include video footage. For instance, such a digital educative feature might show how a teacher supports their students around a challenging task. However, there has been little research on multimedia curriculum, including its use and impact on teacher learning. A recent study began to delve into this void, describing the development, and teachers' use, of multimedia educative curriculum materials (MECMs) focused on supporting teacher learning of scientific argumentation (Loper, McNeill & González-Howard, 2017). These MECMs were made up of videos that included a mixture of classroom footage, interviews with teachers and graphics. Loper and colleagues (2017) found that despite a range in the frequency the MECMs were accessed, teachers were more likely to watch the educative videos when they appeared earlier in the curriculum, and when they aligned to a new activity structure being introduced.

Other work has examined the use of technology more broadly in professional development (PD) contexts. For example, researchers have argued for the importance of online PD – where teachers have an opportunity to view and discuss videos of classroom instruction – as a productive and scalable mechanism by which to support teacher learning, especially given the myriad of responsibilities educators have, and the limited time they have for PD (Dede, Jass

Ketelhut, Whitehouse, Breit, & McCloskey, 2009). Other studies have also shown that teachers can benefit from watching classroom video (Roth, Garnier, Chen, Lemmens, Schwille, & Wickler, 2011). For example, van Es and Sherin (2008) found that teachers engaged in a video club noticed particular aspects of their students' thinking from analyzing and discussing video excerpts of their instruction. Thus, digital technologies hold much promise in transforming and supporting teacher learning (Borko, Whitcomb, & Liston, 2009). Connecting these findings from technology use in PD settings to educative curriculum, multimedia educative materials might better enable teachers to *see* how all students are capable of engaging in the rigorous learning detailed in the NGSS, including science practices. Consequently, in this study we investigated how teachers used multimedia educative curriculum materials and whether that use impacted teachers' beliefs specifically about scientific argumentation.

Methods

In order to address our research question, we conducted a randomized experimental study with both a control group and pretest (Shadish, Cook & Campbell, 2002). Both the control and the experimental groups of teachers taught the same Earth & Space Science middle school curriculum. The one difference was the inclusion of the MECMs. The experimental group teachers received the curriculum incorporating the MECMs, while the control group teachers received the same digital curriculum materials, also delivered through a website, but minus the MECMs. We collected multiple data sources to examine both teachers' use of the curriculum and the impact of the MECM and use on teachers' beliefs and PCK of argumentation. In this section, we first describe the curricular context and participants. Then we describe the data sources and analysis procedures.

Curricular Context

The Earth & Space Science middle school curriculum consisted of 62 lessons organized into three units. In addition, the teachers in the experimental group received the curriculum incorporating the MECMs. The MECMs included 24 videos and 17 interactive reflective prompts to illustrate high quality argumentation in terms of both the structural and dialogic elements.

Participants

To recruit teachers, we sent out e-mails and posted information through the Lawrence Hall of Science network (including our contacts from prior field trials, prior projects (Carnegie, MECM, Middle School, and REESE projects), free advertisements in professional development e-magazines and web pages (National Science Teacher Association, and the California Science Teacher Association), various listervs such as the national GEMS directors network listserv, NASA listserv as well as professional networks from various colleagues, the Lawrence Hall of Science internal staff listserv and the Learning Design Group recruitment database) to solicit interest from middle schools across the country. We accepted the first 98 teacher who expressed interest in enacting the curriculum, completed the pre-survey and permission materials and from whom we received research permission from either the school or school district. In a number of instances, multiple teachers from the same district expressed interest in enacting the curriculum. For these instances, we randomly selected five teachers from the district. We decided on a maximum of five, because that would be 10% of either the MECM group or the control group.

Since districts can have their own initiatives and cultures, we did not want either group to become too impacted by other district variables.

Our study began with 98 teachers – 49 in the MECM treatment group and 49 in the control group. Over the school year, 13 teachers dropped out of the study. We were able to replace some teachers, but stopped in March in order to allow teachers time to enact the three curriculum units. The final group that enacted the curriculum included 90 teachers with 46 teachers in the MECM treatment group and 44 teachers in the control group.

Block randomization. Block randomization to the experimental MECM and control groups occurred at the teacher level in order to control for covariates (Shadish, Cook & Campbell, 2002) that could impact teachers' use and enactment of the curriculum. Specifically, at the school level we considered school type (public, private, charter, faith-based) and school locale (town, city, suburb or rural). At the teacher level, we blocked for years teaching science (0-2, 3-5, 6-10, 11-15, 16+) and highest level of science education (none, bachelors, masters, and doctorate). In addition, teachers were clustered (Ivers et al., 2012) based on school district with all teachers in one district being assigned to the same group. We made this decision since we thought teachers in the same district would potentially talk to each other, which would be problematic if one teacher received the MECM treatment and another teacher received the Control treatment. This could potentially cause unnecessary tensions with a teacher wishing they had received the other group. In addition, a teacher in the MECM treatment could share their resources with a teacher in the Control group. Thirty-eight of the teachers were in a district in which between two and five teachers were enacting the curriculum while sixty teachers were the only participating teacher in their district. Teachers were sorted based on these variables and then we randomly assigned the first teacher (e.g. a coin flip) and alternated after the first assignment clustering all teachers in the same school district.

Clustering can result in the imbalance of covariates across the two groups, which can decrease analytic power and statistical precision as well as the face validity and credibility of a study (Ivers et al., 2012). Consequently, after the random assignment of the teachers, we examined the teacher demographic data from the two groups resulting in two groups with similar school and teacher information (see Table 1).

Table 1: School and Teacher Background Information (n = 90)

	MECM Treatment (n = 46)	**Control (n = 44)**
School Information		
Type of School		
• Public	43	38
• Private	2	2
• Charter	1	2
• Faith-based	0	1
• Other	0	1
School Locale		
• City	16	14
• Suburb	10	13
• Town	6	6
• Rural	14	10
• Not Sure	0	1
Teacher Information		

Years Teaching Science		
• 0-2 years	6	6
• 3-5 years	9	8
• 6-10 years	12	14
• 11-15 years	8	9
• 16+ years	11	7
Highest Level of Science Education		
• No Response	4	1
• Bachelors Degree	38	31
• Masters Degree	3	12
• Doctorate Degree	1	0
Highest Level of Education		
• No Response	1	1
• Bachelors Degree + teaching certificate	26	14
• Masters Degree	18	29
• Doctorate Degree	1	0
Argumentation		
Taught a Curriculum focused on Argument		
• Yes	5	8
• No	41	36
# Workshops, PD or classes on Argument		
• None	27	19
• 1	5	6
• 2 or 3	10	14
• 3 or more	4	5
Included argumentation in Classroom Instruction		
• Never	17	10
• Once	3	1
• A few times	18	19
• Many times	8	14

Teacher background information. Across the two groups, teachers had a wide range of backgrounds including first year teachers to teachers with more than 16 years teaching experience. In addition, teachers had a range of degrees for both science education and education as well as previous experience including argumentation in their classrooms.

Data Collection

We collected multiple data sources: pre surveys, backend curriculum use data (collected through the digital curriculum), self-report curriculum use data (collected through daily lesson surveys), and post surveys. The teacher background information was described previously (See Table ?). We next describe each of the other measures in more detail.

Teacher belief survey. The outcome measure (contained in both the pre and post surveys) included a teacher belief of scientific argument instrument comprising 22 items across three sub-domains – Self-efficacy, Learning Goals, and Student Background and Ability. This instrument was based on a previous pilot, which we then increased the number of items for each construct to improve reliability (McNeill, Katsh-Singer, González-Howard, & Loper, 2016). Principal components analysis was used to confirm the existence of the three sub-scales and Cronbach's alpha was calculated to estimate the pre- and posttest reliabilities. Table 2 shows that the reliability estimates ranged between 0.77 for the posttest Learning Goals sub-scale to 0.93 for the pretest Self-efficacy scale. Sub-scale scores were created by calculating the mean for each teacher across the constituent items to keep the outcome on the same scale (1 to 4) for ease of interpretation.

Table 2: Teacher Beliefs about Scientific Argumentation.

Sub-Scale	Number of Items	Sample Item	Cronbach's Alpha	
			Pretest	Posttest
Teacher Self-Efficacy[a]	8	I feel confident facilitating students' critiques of arguments.	0.93	0.88
Learning Goals[a]	7	Engaging students in argumentation is an important part of learning science.	0.87	0.77
Student Background and Ability[b]	7	Tammy has an IEP for challenges she experiences with reading. Tammy's mother says she likes science and watches TV shows about the environment at home, but is easily frustrated at school.	0.83	0.89

[a] Response options: 1 = Strongly disagree, 2 = Disagree, 3 = Agree, 4 = Strongly Agree
[b] Response options: 1 = Not capable, 2 = Somewhat capable, 3 = Capable, 4 = Very Capable

Backend curriculum use. The curriculum was web-based and as such we were able to access back-end curriculum analytics for each individual teacher. Specifically, we used this data to construct two variables: 1) Number of page views and 2) Number of video views. For all of the teachers (MECM and Control), we were able to see how many pages they viewed over the course of the curriculum enactment. The curriculum consisted of 73 webpages (62 lesson plans and 11 overview and additional resource pages). Consequently, if a teacher looked at each page once their total would be 73, while if they accessed each page three times their total would be 219.

Self-report curriculum use. For each lesson, teachers were asked to complete two short survey questions. The first question asked teachers if they taught the lesson and then provided them with four choices: 1) Yes, I completed the session, 2) Yes, but I only partially completed the session, 3) Yes, but I modified the session, and 4) No, I skipped the session. We used this first item to create two variables. The first variable, Total lessons taught, summed the number of lessons for which a teacher responded 1, 2 or 3. Unfortunately, although there were 62 lessons in the entire unit there were technical difficulties for the surveys inserted into 7 of those lessons. This resulted in missing data for a number of teachers for those lessons. Consequently, we only included the lesson surveys for 55 of the lessons. The second variable calculated the percent of taught lessons that the teacher modified by dividing the number of times they selected choice 3 (Yes, but I modified the lesson), by the total number of lessons. The second lesson survey question asked teachers to check the different parts of the lesson that they used (purpose, preparation, description, right hand notes, videos and powerpoints). We summed these to create two variables one for the total number of non-MECM elements teachers used and the other for the total number of MECM elements used (just for the MECM teachers). Of the seven lessons that were removed because of technical difficulties, only two of those lessons included MECMs.

In addition, at the end of the enactment, teachers were asked to reflect on how frequently they used the different curricular resources. Specfically, they were asked – *How frequently did you use the following teacher supports about argumentation?*

Data Analysis

First, we conducted dependent means t-tests to examine whether there were any significant pre-posttest differences across all teachers. Next we examined descriptives to determine whether there was variation in teachers' use of the materials. Finally, we used multiple linear regression to examine the impact of the MECM treatment and curriculum use on teachers' beliefs. Specifically, teachers' posttest beliefs were regressed on their pretest beliefs, an indicator of their group membership, and a series of background characteristics and curriculum indicators. A power analysis was conducted using G*Power to estimate the minimum effect size detectable with a linear regression model given a fixed sample size, power, and significance level. With 90 teachers, power = 0.80, and α = 0.05, linear regression model with a maximum of 4 predictors will allow us to detect a medium effect size of $f^2 = 0.15$. Given the relatively limited sample size and statistical power, the models were formulated sequentially to include blocks of variables. For each outcome, a parsimonious final model that included only group membership and any statistically significant predictors from previous blocks was formulated.

Results

Overall Changes in Teacher Beliefs

Table 3 presents the mean pre- and posttest scores for the three sub-scales, along with the results of dependent means t-tests. The results indicate that all teachers' posttest beliefs were significantly higher than teachers' pretest belief scores regardless of which version of the curriculum that they enacted. This suggests that teaching a curriculum with an explicit focus on argumentation supported changes in teachers' beliefs in terms of self-efficacy, learning goals and student background. For self-efficacy, they became more confident in their own ability to support students in argumentation, such as facilitating students critique of arguments. The increase in learning goals suggests that teachers felt argumentation was even more important for science instruction, for goals such as increasing students' reasoning skills. Finally, after teaching the unit, teachers' beliefs that different students are capable of engaging in argumentation increased.

Table 3: Changes in Teacher's Beliefs about Argumentation

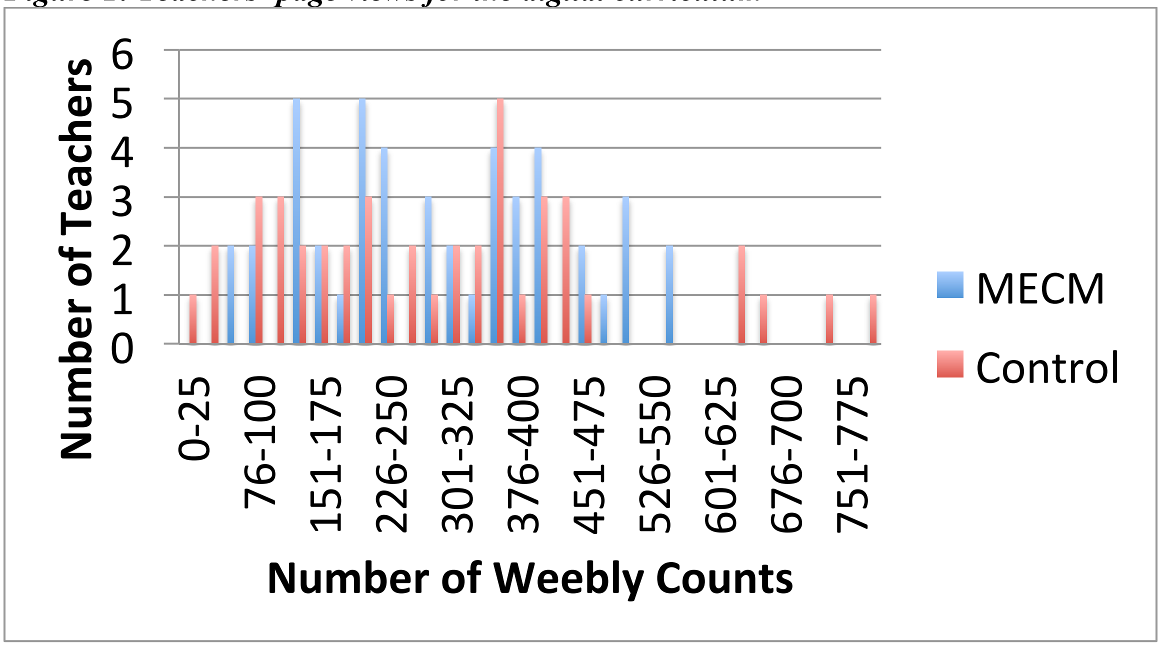
	t	Pretest		Posttest		Paired Samples t-test	
	N	Mean (se)	SD	Mean (se)	SD		Sig.†
Self Efficacy	90	3.09(.05)	0.52	3.51(.04)	0.38	-8.31	<.001
Learning Goals	90	3.60(.04)	0.39	3.82(.04)	0.25	-5.59	<.001
Student Background	90	3.02(.05)	0.43	3.31(.06)	0.54	-5.22	<.001

† A Bonferonni adjustment was applied to account for multiple tests (α/3 = 0.017)

Variation in Teacher's Use of the Curriculum

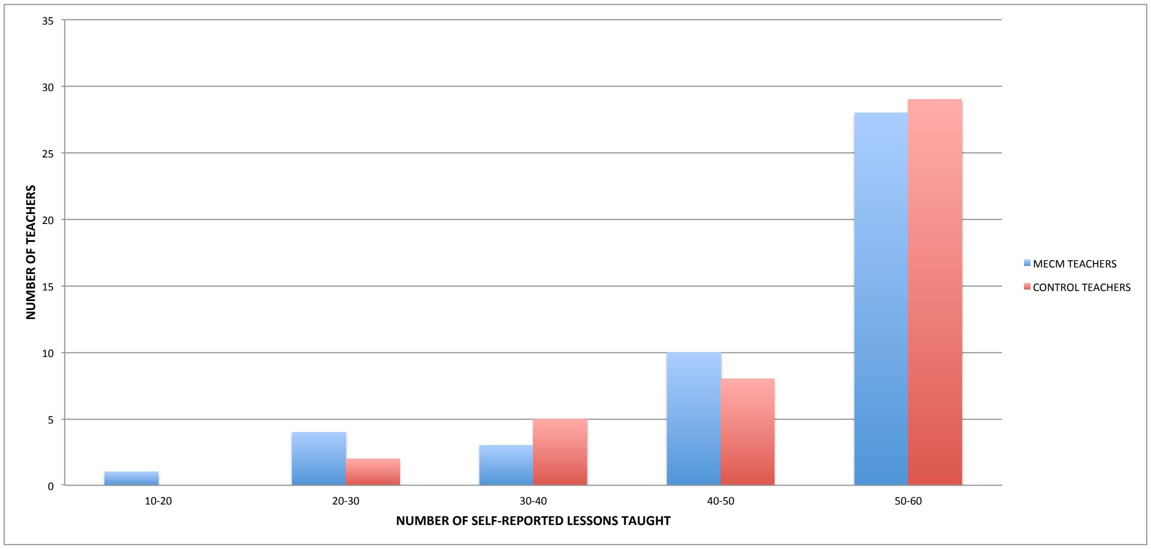
In terms of teachers' use of the curriculum, we looked at both *Backend Curriculum Use* data (collected automatically by the digital curriculum) and *Self-report Curriculum Use* data (collected through teacher surveys within each lesson and a post survey). Across all variables, we saw a wide range in teacher use. For example, the *Backend Curriculum Use* data showed that teachers viewed the pages of the digital curriculum between 14 and 785 times, with a mean of 301. There was similar use across the MECM and Control groups, with the average for MECM at 299 page views and the average for Control at 304 page views (see Figure 1).

Figure 1: Teachers' page views for the digital curriculum



The *Self-report Curriculum Use* data included that teachers reported teaching between 11 and 55 lessons, with a mean of 48 lessons. Again, the data was fairly consistent across the MECM and Control groups. The MECM group reported teaching an average of 47 lessons while the Control group reported teaching an average of 49 lessons out of a total possible of 62 lessons.

Figure 2: Teachers' self-reported number of lessons taught



Model for Self-Efficacy

Table 4 includes the self-efficacy regression models. Model 1 shows that the only *Teacher Background* variable associated with teachers' post self-efficacy beliefs was their pre self-efficacy beliefs. Teacher who were more confident in their abilities to teach argumentation before the enactment were more confident after the enactment. After controlling for the effects of any differences in teachers' backgrounds, we next entered in the *Condition* variable – Control group versus MECM treatment (Model 2). Just entered by itself, *Condition* was not significant. In Model 3, we added the *Backend Curriculum Use* variables. Neither the total page views nor the total videos views (for the MECM group) were associated with teachers' posttest self-efficacy for argumentation. Finally, we entered teachers' *Self-Report Curriculum Use* variables (Model 4). As teachers reported teaching more lessons, their post self-efficacy increased. For example, on average, a teacher who taught all of the lessons would have a 0.22 higher score than a teacher who taught half the lessons on a scale of 1 = Strongly disagree to 4 = Strongly Agree. With the addition of the *Self-Report* variables, the *Condition* also becomes significant with MECM teachers having lower post self-efficacy compared to the Control teachers.

Model for Learning Goals

Table 5 includes the learning goals regression models. Similar to self-efficacy, the only significant teacher background variable was the teachers' pre-survey score for learning goals. Neither the condition, nor the backend curriculum use data had a significant impact on teachers' post beliefs about the importance of argumentation learning goals. In fact the only significant predictor in this model, was the number of lessons the teachers reported teaching. The more lessons a teacher taught, the more their post learning goals score increased. For example, on average a teacher who taught all the lessons would have a 0.19 higher score than a teacher who taught half the lessons. This suggests that the number of the lessons taught impacts any changes in teachers' beliefs.

Model for Student Background

Table 6 includes the student background regression models. Again, similar to the last two models, the only significant teacher background variable was teachers' pre-survey score for student background. In this model, neither Condition nor Self-report Curriculum Use had a significant effect on teachers' changes in students' abilities to engage in argumentation. The only significant variable in this model is the number of videos that the teachers watched. Teachers who watch mode videos, which was only available in the MECM treatment, had lower post student background scores compared to the Control teachers. This means that on average they rated the descriptions of students with different background as less capable of engaging in argumentation.

Table 4: Teacher Self Efficacy Regression Models for Scientific Argumentation

	Model 1 Demographics	Model 2 Condition	Model 3 Backend Use	Model 4 Self-Report Use
	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)
Intercept	2.44(0.24)***	2.50(0.22)***	2.38(0.23)***	2.15(0.33)***
Teacher Background				
Pre Belief Survey	0.36(0.07)***	0.35(0.07)***	0.36(0.07)***	0.35(0.07)***
Years Teaching	-0.02(0.03)			
Degree	0.08(0.08)			
Race	-0.13(0.10)			
Gender	-0.01(0.07)			
Condition				
Group		-0.13(0.07)	-0.09(0.08)	-0.14(0.07)*
Backend Curriculum Use				
Total Page Views†			<0.01(<0.01)	
Total Video Views			-0.01(0.01)	
Self-report Curriculum Use				
# Lessons Taught†				0.07(0.03)*
Percentage Modified				-0.23(0.14)
Comfort with Digital				-0.02(0.05)
Way Used Curriculum				0.03(0.04)
Adjusted R^2	22.0%	24.9%	25.8%	29.9%

**** Significant at p < .001, ** Significant at p < .01, * Significant at p < .05
† Variable rescaled to represent a change of 10 units.

Table 5: Learning Goals Regression Models for Scientific Argumentation

	Model 1 Demographics	Model 2 Condition	Model 3 Backend Use	Model 4 Self-Report Use
	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)
Intercept	2.89(0.24)***	2.94(0.23)***	2.91(0.23)***	2.69(0.29)***
Teacher Background				
Pre Belief Survey	0.25(0.06) ***	0.24(0.06)***	0.24(0.06)***	0.23(0.06)***
Years Teaching	0.02(0.02)			
Degree	-0.05(0.05)			
Race	0.01(0.07)			
Gender	<0.01(0.05)			
Condition				
Group		0.02(0.05)	0.01(0.06)	0.02(0.05)
Backend Curriculum Use				
Total Page Views†			<0.01(<0.01)	
Total Video Views			<0.01(0.01)	
Self-report Curriculum Use				
# Lessons Taught†				0.06(<0.02)*
Percentage Modified				-0.14(0.10)
Comfort with Digital				<0.01(0.03)
Way Used Curriculum				0.02(0.03)
Adjusted R^2	12.7%	12.8%	12.5%	18.7%

**** Significant at p < .001, ** Significant at p < .01, * Significant at p < .05

† Variable rescaled to represent a change of 10 units.

Table 6: Learning Goals Regression Models for Scientific Argumentation

	Model 1 Demographics	Model 2 Condition	Model 3 Backend Use	Model 4 Self-Report Use
	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)	Coeff (s.e.)
Intercept	1.53(0.39)***	1.59(0.36)***	1.52(0.36)***	1.33(0.55)*
Teacher Background				
Pre Belief Survey	0.59(0.12)***	0.57(0.12)***	0.58(0.12)***	0.59(0.12)***
Years Teaching	-0.03(0.40)			
Degree	0.20(0.11)			
Race	-0.14(0.14)			
Gender	0.02(0.10)			
Condition				
Group		-0.02(0.10)	0.13(0.12)	0.11(0.12)
Backend Curriculum Use				
Total Page Views†			<0.01(<0.01)	
Total Video Views			-0.02(0.01)*	-0.02(0.01)
Self-report Curriculum Use				
# Lessons Taught†				0.09(0.05)
Percentage Modified				-0.02(0.22)
Comfort with Digital				-0.02(0.07)
Way Used Curriculum				-0.06(0.06)
Adjusted R^2	21.1%	19.1%	21.8%	22.9%

**** Significant at p < .001, ** Significant at p < .01, * Significant at p < .05
† Variable rescaled to represent a change of 10 units.

Discussion

Changing teachers' beliefs about science instruction can be a difficult and time consuming task (Bryan, 2012). The results from this study suggest that enacting a digital curriculum with a specific focus on argumentation can support positive changes in teachers' beliefs about argumentation. Across all three sub-scales, we observed significant growth in teachers' beliefs. This suggests that enacting a curriculum with a specific focus on argumentation can increase teachers' self-efficacy, views about the importance of argumentation learning goals, and views of how capable students are with diverse backgrounds.

We also saw a wide range in how teachers used the digital curriculum. For self-efficacy, this differential use did impact teacher's beliefs. Specifically, the number of lessons teachers reported enacting with their students impacted their beliefs about argumentation. Teachers became more confident in their ability to teach argumentation and believed argumentation was more important for science instruction as they enacted more lessons. This suggests the importance of teacher learning experiences around argumentation being directly connected to their classroom instruction. The act of trying argumentation lessons with their students can have a positive impact on teachers' beliefs (Knight & McNeill, 2016)

The MECM condition impacted teachers' beliefs for self-efficacy; however, this impact was negative. The increase in self-efficacy was lower for the teachers receiving the MECM treatment, which included videos of students engaged in argumentation. One potential reason for this is the videos may have *problematized* (Reiser, 2004) what teachers thought counted as argumentation, as well as potentially their role as a teacher in a classroom focused on this science practice. The videos may have altered teachers' views of argumentation to be more complicated then they envisioned at the begging of the curriculum enactment. Observing expert teachers engage in argumentation with their students, may have resulted in smaller increases in self-efficacy because the ultimate goal they now envisioned was different.

Furthermore, the number of MECM videos watched impacted teachers' beliefs about students' abilities. The more videos teachers watched, on average the less growth occurred in their beliefs in this area. These survey items explicitly described students with a range of backgrounds, such as students on IEPs and English Language Learners. In retrospect, the videos did not explicitly discuss diverse students or purposefully highlight students with some of the backgrounds described in the items. Future MECMs should try and incorporate a more diverse range of students, to better highlight all students' abilities to engage in argumentation.

Multimedia images of argumentation may provide different images of this practice compared to both text-only curriculum materials and teachers' prior conceptions. These images can potentially change what teachers think counts as argumentation. Furthermore, as the field engages more in developing digital resources, we need to consider the range of students and classrooms depicted in these resources.

References

Alozie, N. M., Moje, E. B., & Krajcik, J. S. (2010). An analysis of the supports and constraints for scientific discussion in high school project-based science. *Science Education*, 94(3), 395–427.

Arias, A. M., Bismack, A. S., Davis, A. E. & Sullivan Palincsar, A. (2016). Interacting with a suite of educative features: Elementary science teachers' use of educative curriculum materials. *Journal of Research in Science Teaching, 53*(3), 422-449.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6–8, 14.

Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education*, 95, 191–216

Borko, H., Whitcomb, J., & Liston, D. (2009). Wicked problems and other thoughts on issues of technology and teacher learning. *Journal of Teacher Education*, *60*(1), 3-7.

Bryan, L. A. (2012). Research on science teacher beliefs. In B. J. Fraser, K. Tobin and C. J. McRobbie (Eds.), *Second international handbook of science education*. (Vol. 1. pp. 477-495). Dordrecht: Springer.

Davis, E. A., Janssen, F. J., & Van Driel, J. H. (2016). Teachers and science curriculum materials: where we are and where we need to go. *Studies in Science Education*, *52*(2), 127-160.

Davis, E.A. & Krajcik, J.S. (2005) Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34* (3), 3 -14.

Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A research agenda for online teacher professional development. *Journal of teacher education*, *60*(1), 8-19.

Duncan, R. G., El-Moslimany, H., McDonnell, J., & Lichtenwalner, S. (2011). Supporting Teachers' Use of a Project-Based Learning Environment in Ocean Science: Web-Based Educative Curriculum Materials. *Journal of Technology and Teacher Education*, *19*(4), 449-472.

Evagorou, M. & Dillon, J. (2011). Argumentation in the teaching of science. In Corrigan, D., Dillon, J. & Gunstone, R. (Eds.), *The professional knowledge base of science teaching.* (pp. 189-204). New York: Springer.

Ford, M. J. (2012). A dialogic account of sensemaking in scientific argumentation and reasoning. *Cognition and Instruction*, *30*(3), 207-245.

Ivers, N. M., Halpern, I. J., Barnsley, J., Grimshaw, J. M., Shah, B., R., Tu, K., Upshur, R. & Zwarenstein, M. (2012). Allocation techniques for balance at baseline in cluster randomized trials: A methodological review. *Trials*, *13*(120), 1-9.

Jiménez-Aleixandre, M. P. & Erduran, S. (2008). Argumentation in science education: An overview. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.). *Argumentation in science education: Perspectives from classroom-based research.* (pp. 3-28), Dordrecht: Springer.

Jones, M. G. & Leagon, M. (2014). Science teacher attitudes and beliefs: Reforming practice. In Lederman, N. G., & Abell, S. K. (Eds.), *Handbook of research on science education* (Vol. 2), 830-847. NY: Routledge.

Katsh-Singer, R., McNeill, K. L., & Loper, S. (2016). Scientific Argumentation for All? Comparing Teacher Beliefs About Argumentation in High, Mid, and Low Socioeconomic Status Schools. *Science Education*, 100(3), 410-436.

Knight-Bardsley, A. M. & McNeill, K. L. (2016). Teachers' pedagogical design capacity for scientific argumentation. *Science Education. 100*(4), 645-672.

Krajcik, J., & Delen, I. (2017). The Benefits and Limitations of Educative Curriculum Materials. *Journal of Science Teacher Education*, *28*(1), 1-10.

Loper, S., McNeill, K. L. & González-Howard, M. (2017). Multimedia educative curriculum materials (MECMs): Teachers' use of MECMs to support argumentation. *Journal of Science Teacher Education, 28*(1), 36-56.

Marco-Bujosa, L. M., McNeill, K. L., González-Howard, M., & Loper, S. (2017). An exploration of teacher learning from an educative reform-oriented science curriculum: Case studies of teacher curriculum use. *Journal of Research in Science Teaching*, *54*(2), 141-168.

McNeill, K. L. (2009). Teachers' use of curriculum to support students in writing scientific arguments to explain phenomena. *Science Education, 93*(2), 233–268.

McNeill, K. L., González-Howard, M., Katsh-Singer, R., & Loper, S. (2016). Pedagogical content knowledge of argumentation: Using classroom contexts to assess high-quality PCK rather than pseudoargumentation. *Journal of Research in Science Teaching,* 53(2), 261-290.

McNeill, K. L., Katsh-Singer, R., González-Howard, M., & Loper, S. (2016). Factors impacting teachers' argumentation instruction in their science classrooms. *International Journal of Science Education, 38*(12), 2026-2046.

McNeill, K. L., Lizotte, D. J, Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153-191.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States (Appendix F).* Washington, DC: The National Academies Press.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, 328, 463-466.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994–1020

Pruitt, S. L. (2014). The next generation science standards: The features and challenges*. Journal of Science Teacher Education, 25*(2), 145-156.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences. 13*(3), 273-304.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, *75*(2), 211-246.

Roth, K.J., Garnier, H.E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N.I.Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*. 48(2), 117-148.

Sampson, V., & Blanchard, M. R. (2012). Science teachers and scientific argumentation: Trends in views and practice. *Journal of Research in Science Teaching,* 49(9), 1122-1148.

Schneider, R. M., Krajcik, J. & Blumenfeld, B. (2005). Enacting reform-based science materials: Range of teacher enactments in reform classrooms. *Journal of Research in Science Teaching, 42*(3), 283-312.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin Company.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and teacher education*, *24*(2), 244-276.

Zohar, A. (2008). Science teacher education and professional development in argumentation. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 245–268), Dordrecht: Springer.